

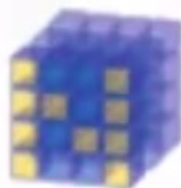
# Scientifics Computing Libraries in Python

## 1. Scientifics Computing Libraries



### Pandas

(Data structures & tools)



### NumPy

(Arrays & matrices)



# Visualization Libraries in Python

## 2. Visualization Libraries



### **Matplotlib**

(plots & graphs, most popular)



### **Seaborn**

(plots : heat maps, time series, violin plots)

Another high-level visualization library,  
Seaborn, is based on matplotlib.

# Machine Learning and Deep Learning Libraries in Python

3. Machine Learning and Deep Learning



## Scikit-learn

(Machine Learning : regression, classification, ...)



Keras

## Keras

(Deep Learning Neural Networks, ...)

For deep learning, Keras enables you to build the standard deep learning model.

# Deep Learning Libraries in Python

3. Deep Learning Libraries



**TensorFlow**

(Deep Learning: Production and Deployment)



**PyTorch**

(Deep Learning: regression, classification,... )

Pytorch is used for experimentation, making it simple for researchers to test their ideas

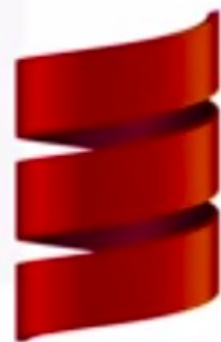
# Scala-Libraries

---

- Vegas
- Deep Learning: Big DL

**VEGAS**

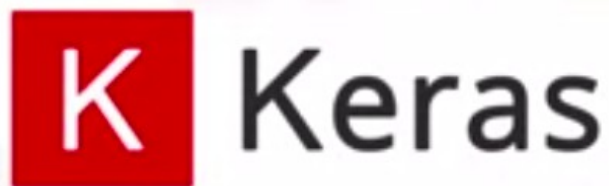
**BigDL**



# R- Libraries

---

- Ggplot2
- You can also Libraries that allow you to interface with Keras and TensorFlow





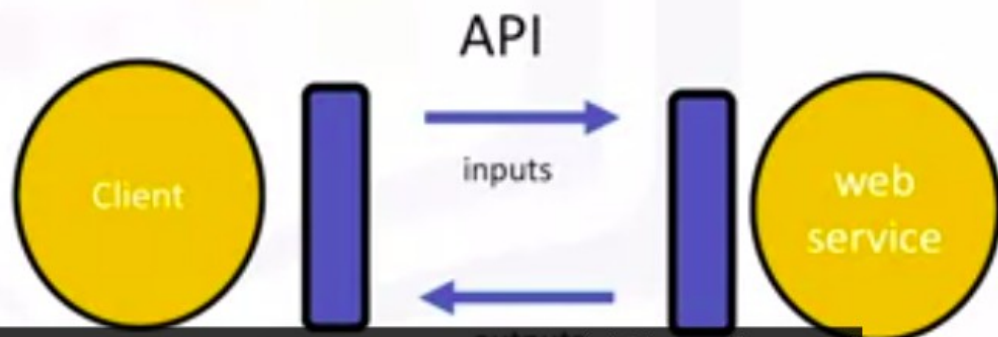
# REST APIs

**RE**presentational **S**tate **T**ransfer APIs

“Representational,” the S stands for “State,”  
the T stand for “Transfer.” In rest APIs,

# REST APIs

- REST APIs are used to interact with web services, i.e., Applications that you call through the internet
- They have a set of Rules regarding:
  1. Communication
  2. Input or Request
  3. Output or Response



Output or Response. Here are some common API-related terms. You or your code can be thought of





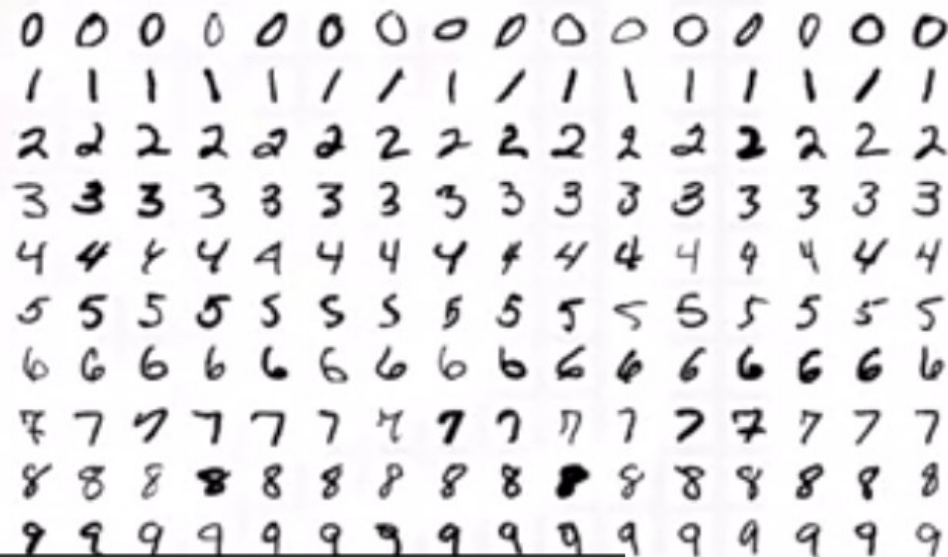
the audio file to the API; this process is called a post request. The API then sends



The Watson Language-Translator API provides another example. You send the text you would

# What's a data set

- Collection of data
- Data structures
  - Tabular data
  - Hierarchical data, network data
  - Raw files



It contains images of handwritten digits and is commonly used to train image processing

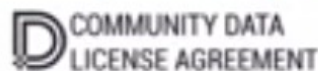
# Data Ownership

---

- Private data
  - Confidential
  - Private or personal information
  - Commercially sensitive
- Open data
  - Scientific institutions
  - Governments
  - Organizations
  - Companies
  - Publicly available



# Community Data License Agreement



[Home](#) [FAQ](#) [CDLA Versions](#) [Reference Translations](#) [Join the Discussion](#) [Q](#)

## COMMUNITY DATA LICENSE AGREEMENT

Collaborative licenses to enable access, sharing and use of data openly among individuals and organizations

- <http://cdla.io> – A Linux Foundation project
- CDLA-Sharing: Permission to use and modify data; publication only under same terms
- CDLA-Permissive: Permission to use and modify data; no obligations



# The Data Asset eXchange

## Data Asset eXchange

Explore useful and relevant data sets for enterprise data science

Dataset | JSON Lines

MedNLI

Sep 17, 2019



Dataset | CSV

Fashion-MNIST

Sep 12, 2019



Dataset | CSV

Contract Bank

Sep 12, 2019

Dataset | CSV

NOAA Weather Data -  
JFK Airport

Dataset | CSV, H264

Double Pendulum  
Chaotic

Dataset | CSV

Nutch

- Curated collection of data sets:
  - From IBM Research and 3<sup>rd</sup> party
  - Multiple application domains
- Data Science friendly licenses

<https://developer.ibm.com/exchanges/data/>



# What is a model?

---

- Data can contain a wealth of information
- Machine Learning (ML) models identify patterns in data
- A model must be trained on data before it can be used to make predictions
- Supervised, unsupervised and reinforcement learning are types of ML

# Supervised Learning

---

- Supervised Learning
  - Data is labeled and model trained to make correct predictions
- Regression
  - Predict real numerical values
  - e.g. home sales prices, stock market prices
- Classification
  - Classify things into categories
  - e.g. email spam filters, fraud detection, image classification

# Other learning types

---

- Unsupervised Learning
  - Data is not labeled
  - Model tries to identify patterns without external help
  - Common learning problems: clustering and anomaly detection
- Reinforcement Learning
  - Conceptually similar to human learning processes
  - e.g. a robot learning to walk; chess, Go and other games of skill

# Deep Learning

---

- Tries to loosely emulate how the human brain works
- Applications
  - Natural Language Processing
  - Image, audio, and video analysis
  - Time series forecasting
  - Much more
- Requires typically very large datasets of labeled data and is compute intensive

# Deep Learning Models

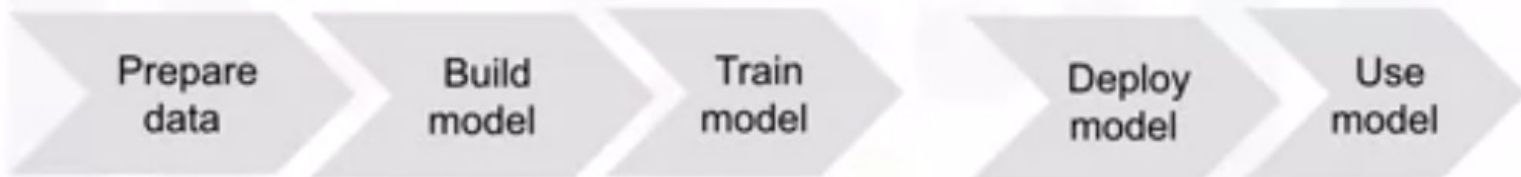
---

- Build from scratch or download from public model repositories
- Built using frameworks, such as:
  - TensorFlow
  - PyTorch
  - Keras
- Popular model repositories
  - Most frameworks provides a “model zoo”
  - ONNX model zoo



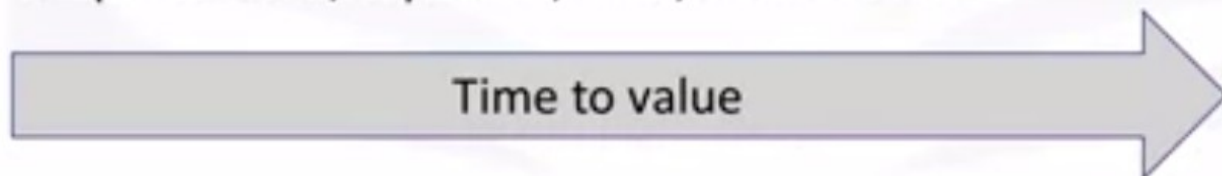
# Using models to solve a problem

What is this?



Iterative process:

Requires data, expertise, time, and resources



This is a teddy bear



# Summary

<https://developer.ibm.com/exchanges/models/>

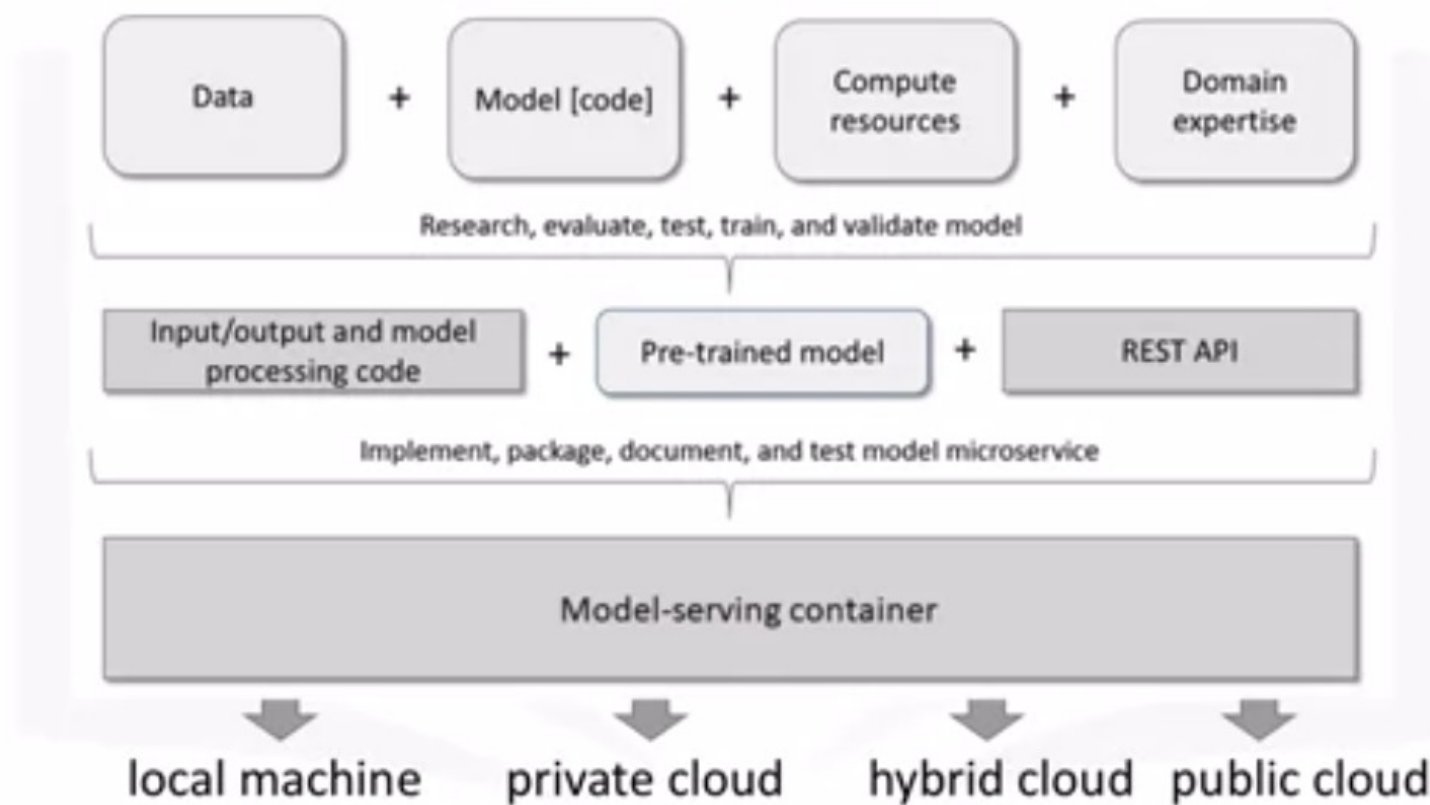
The screenshot displays the IBM Model Asset eXchange interface. At the top, a dark banner reads "Model Asset eXchange" with the tagline "Free, deployable, and trainable code. A place for developers to find and use free and open source deep learning models." Below this, there are tabs for "Featured", "Deployable", and "Trainable". Two featured models are shown: "Toxic Comment Classifier" (deployable, dated Jun 04, 2019) and "Text Sentiment Classifier" (deployable and trainable, dated Mar 25, 2019). An overlay window provides a detailed view of the "Object Detector" model, which is deployable and trainable. It includes a description: "Localize and identify multiple objects in a single image." and offers links to "Try the tutorial", "Join the community", "Try the API", "Try the web app", and "Try in a Node-RED flow". The "Overview" section for the Object Detector model states it recognizes objects from 80 classes in the COCO dataset, uses a deep convolutional neural network for feature extraction, and was trained on the COCO dataset. It is based on the "COCO Multiscale FC object detection model for TensorFlow".

# MAX reduces time to value

---

- Free open-source deep learning microservices
  - Use pre-trained or custom-trainable state-of-the-art models
  - Fully tested, deploy in minutes
  - Approved for personal and commercial use
- Available for variety of domains:
  - Object detection (“which objects are in this image”)
  - Image, audio, and text classification (“what is in this ...”)
  - Named Entity recognition (“identify entities in text”)
  - Image-to-Text translation (“generate image caption”)
  - Human pose detection

# MAX model-serving microservice



# Model-serving microservice API

- Model-serving microservices expose standardized REST API

## MAX Object Detector <sup>1.2.0</sup>

**model** Model information and inference operations

**POST**

**/model/predict** Make a prediction given input data

**GET**

**/model/labels** Return the list of labels that can be predicted by the model

**GET**

**/model/metadata** Return the metadata associated with the model

Prediction/scoring endpoint

Metadata endpoints

- Application-friendly inputs and outputs

# Prediction request handling

